Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 165 (2016) 257–262

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Notes

Fairness and group-strategyproofness clash in assignment problems

Sophie Bade [a,b,*,1]

[a] *Royal Holloway College, University of London, United Kingdom*
[b] *Max Planck Institute for Research on Collective Goods, Bonn, Germany*

Received 15 June 2015; final version received 18 April 2016; accepted 19 April 2016

Available online 22 April 2016

Abstract

No group-strategyproof and ex-post Pareto optimal random matching mechanism treats equals equally. Every mechanism that arises out of the randomization over a set of non-bossy and strategyproof mechanisms is non-bossy. Random serial dictatorship, which arises out of a randomization over all deterministic serial dictatorships is non-bossy but not group-strategyproof.

JEL classification: C78

Keywords: Random matching mechanism; Non-bossy; Group-strategyproof; Random serial dictatorship

1. Introduction

An ideal mechanism would be impossible to manipulate, efficient and fair. For house matching problems, where a finite set of agents with linear preferences over houses needs to be matched to these houses, there exists no group-strategyproof and ex-post Pareto optimal random matching mechanism that treats equals equally. So group-strategyproofness, one of the strongest non-

* Correspondence to: Royal Holloway College, University of London, United Kingdom
E-mail address: sophie.bade@rhul.ac.uk.

1 Support from the ARCHES prize is gratefully acknowledged. I would like to thank Anna Bogomolnaia, Annika Johnson, Debasis Mishra and Herve Moulin.

http://dx.doi.org/10.1016/j.jet.2016.04.012

manipulability requirements, clashes with some of the weakest efficiency- and fairness-criteria: ex-post Pareto optimality and equal treatment of equals. A mechanism that maps every profile of preferences to a lottery over Pareto optima is ex-post Pareto optimal. The mechanism treats equals equally, if any two agents who submit the same preference face the same lottery over houses. It is group-strategyproof if no group ever gains by lying about the group's preferences. Group-strategyproofness strengthens strategyproofness which only requires that no agent ever gains by lying about his preferences.

My result complements Bogomolnaia and Moulin's (2001) theorem that no strategyproof and ordinally efficient matching mechanism treats equals equally. While Bogomolnaia and Moulin and I use the same weak criterion of fairness, our efficiency and non-manipulability requirements differ: where I only impose ex-post Pareto optimality, Bogomolnaia and Moulin require ordinal efficiency; where they only impose strategyproofness, I require group-strategyproofness.

Papai (2000) has given a very useful characterization of group-strategyproofness: a deterministic matching mechanism is group-strategyproof if and only if it is strategyproof and non-bossy, in the sense that no agent can change another agent's outcome without also changing his own. While Barbera, Berga, and Moreno (2014) show that this equivalence extends far beyond house-matching problems, Thompson (2016) provides an overview on where the equivalence holds (and where it fails). I apply Theorem 1 to the question whether this equivalence carries over to random matching mechanisms. Random serial dictatorship, which arises out of a uniform randomization over the order of agents as dictators in serial dictatorship, is ex-post Pareto optimal and treats equals equally. By Theorem 1, random serial dictatorship is not group-strategyproof. To see that Papai's (2000) equivalence result does not extend to the random matching context it only remains to show that random serial dictatorship, which is known to be strategyproof, is non-bossy. To this end, Theorem 2 shows that any randomization over a set of non-bossy and strategyproof mechanisms yields a non-bossy random matching mechanism. Since any serial dictatorship is non-bossy and strategyproof, random serial dictatorship is non-bossy.

The results presented here cover matching problems with and without outside options. In fact, a few minor changes suffice to extend the proofs from one case to the other. This contrasts with other results in matching that crucially depend on the presence (or absence) of outside options. Svensson's (1999) characterization of serial dictatorship does not extend to the case with outside options. Kesten and Kurino (2013) show that while deferred acceptance is an optimal mechanism when considering the full domain of preferences with outside options, improvements upon deferred acceptance are possible if (at least some) agents have no outside options. Erdil (2014) constructs a random matching mechanism that ex-ante Pareto dominates random serial dictatorship when agents have outside options.

2. Definitions

There is a set $N = \{1, \ldots, n\}$ of agents and a set of houses H. The option to stay homeless ($\emptyset$) is always available $\emptyset \in H$. Generic elements of H (including $\emptyset$) are denoted h. A matching is a set of agent-house pairs, denoted as a vector $x \in H^n$ where $x_i = x_j$ and $i \neq j$ imply $x_i = \emptyset$. Under x agent i is unmatched if $x_i = \emptyset$, otherwise house x_i is agent i's is match. The set of all matchings is X. Agent i's preference on H is a linear order $\succsim_i$. So $\succsim_i$ is complete, transitive and $h \sim h'$ implies $h = h'$. A profile of all agents' preferences $(\succsim_i)_{i \in N}$ is denoted $\succsim$, where $\succsim_G$ and $\succsim_{-G}$ are the preferences of all agents in some group $G \subset N$ and outside that group, so $\succsim = (\succsim_G, \succsim_{-G})$. The set Ω is the set of all profiles of linear orders on H. Agents are selfish in the sense that they only consider their own houses when ranking different matchings.

The grand set of matching problems, described by Ω and X allows for outside options: matchings may leave some agents and houses unmatched and agents may prefer homelessness to some houses ($\emptyset \succ_i h$ holds for some $\succsim \in \Omega$, $i \in N$ and $h \in H$). Throughout the text $\emptyset \neq \tilde{\Omega} \subset \Omega$ stands for an arbitrary, non-empty domain of linear orders, with $(\succsim'_G, \succsim_{-G}) \in \tilde{\Omega}$ for any $\succsim, \succsim' \in \tilde{\Omega}$ and $G \subset N$. The domain for which all agents rank homelessness $\emptyset$ as their worst option is denoted $\hat{\Omega}$. The subset $\hat{X} \subset X$ of matchings without outside options, is such that the number of unmatched agents under any $x \in \hat{X}$ is minimal. If there are enough houses for all agents to be matched ($| H | \geq n$) then x is a matching without outside options if and only if there is no agent $i \in N$ with $x_i = \emptyset$. Together $\hat{\Omega}$ and $\hat{X}$ describe the set of matching problems without outside options.

A **mechanism** ϕ maps $\tilde{\Omega}$ to the set of matchings X and agent i obtains $\phi_i(\succsim)$ under ϕ at the profile $\succsim$. A **random mechanism** ρ maps $\tilde{\Omega}$ to ΔX, where the set of all lotteries on some finite S is denoted ΔS. Under ρ agent i faces the lottery $\rho_i(\succsim) \in \Delta H$ at $\succsim$ where $\rho_i(\succsim)(h) := \rho(\succsim)(\{x \mid x_i = h\})$ is the probability that i is matched to h under $\rho(\succsim)$.

The mechanism $\rho : \tilde{\Omega} \to \Delta X$ is **(ordinally) group-strategyproof** if any group-deviation that changes the outcomes for this group, renders some group-member ordinally worse off, in the sense that his probability to receive a house better than h' weakly decreases for all h' and strictly decreases for some h^*. So ρ is group-strategyproof if for all $(\succsim, \succsim'_G, G)$ with $\rho_i(\succsim) \neq \rho_i(\succsim'_G, \succsim_{-G})$ for some $i \in G$, there exist some $i^* \in G$ and $h^* \in H$ such that $\sum_{h \succsim_{i^*} h^*} \rho_{i^*}(\succsim)(h) > \sum_{h \succsim_{i^*} h^*} \rho_{i^*}(\succsim'_G, \succsim_{-G})(h)$ and $\sum_{h \succsim_{i^*} h'} \rho_{i^*}(\succsim)(h) \geq \sum_{h \succsim_{i^*} h'} \rho_{i^*}(\succsim'_G, \succsim_{-G})(h)$ for all $h' \in H$. Restricting attention to singleton groups G, (ordinal) group-strategyproofness reduces to **(ordinal) strategyproofness** and ρ is ordinally strategyproof if $\sum_{h \succsim_i h'} \rho_i(\succsim)(h) \geq \sum_{h \succsim_i h'} \rho_i(\succsim'_i, \succsim_{-i})(h)$ holds for all $(\succsim, i, \succsim'_i)$ and $h' \in H$. The mechanism $\rho : \tilde{\Omega} \to \Delta X$ satisfies **equal treatment of equals** if any two agents who announce the same preferences face the same distribution, so $\rho_i(\succsim) = \rho_j(\succsim)$ if $\succsim_i = \succsim_j$. A mechanism ρ is **ex-post Pareto optimal** if $\rho(\succsim)(x) > 0$ implies that x is Pareto optimal at $\succsim$.

3. Group-strategyproofness

Group-strategyproofness clashes with even the mildest criteria of fairness and efficiency.

Theorem 1. *If there are at least three agents and three houses, then no ex-post Pareto optimal, group-strategyproof mechanism $\rho : \Omega \to \Delta X$ treats equals equally.*

Proof. Suppose the ex-post Pareto optimal and group-strategyproof mechanism $\rho : \Omega \to \Delta X$ did treat equals equally. Let $\{a, b, c\} \subset H \setminus \{\emptyset\}$. Say the preferences $\succsim_i$, $\succsim'_1$, and $\succsim^\circ_2$ all rank a, b, and c above any other house. The three preferences agree on all houses other than a, b, and c (so $\succsim_i$, $\succsim'_1$, and $\succsim^\circ_2$ coincide on $H \setminus \{a, b, c\}$). Their different rankings of a, b, and c are given by the following table

$\succsim_i$	a	b	c
$\succsim'_1$	b	a	c
$\succsim^\circ_2$	a	c	b

Let $\succsim$ be such that each agent has the preference $\succsim_i$. At $\succsim$, $(\succsim'_1, \succsim_{-1})$, and $(\succsim^\circ_2, \succsim_{-2})$ all $n \geq 3$ agents prefer a, b and c to all other houses (and to homelessness) and a, b and c must, by the ex-post Pareto optimality of ρ, be matched with probability 1 under $\rho(\succsim)$, $\rho(\succsim'_1, \succsim_{-1})$, and

$\rho(\succsim_2^\circ, \succsim_{-2})$. By equal treatment of equals, each agent obtains a, b, c with probability $\frac{1}{n}$ under $\rho(\succsim)$. Since ρ is ex-post Pareto optimal, agent 1 never gets a under $\rho(\succsim_1', \succsim_{-1})$. Equal treatment of equals then implies that each agent $i \neq 1$ receives a with probability $\frac{1}{n-1}$ under $\rho(\succsim_1', \succsim_{-1})$. Since ρ is strategyproof, agent 2 must obtain a with probability $\frac{1}{n}$ under $\rho(\succsim_2^\circ, \succsim_{-2})$. By equal treatment of equals, all other agents must equally share the remaining probability mass $\frac{n-1}{n}$, implying $\rho_i(\succsim_2^\circ, \succsim_{-2})(a) = \frac{1}{n}$ for all $i \in N$. Since ρ is ex-post Pareto optimal agent 2 never gets b under $\rho(\succsim_2^\circ, \succsim_{-2})$. Equal treatment of equals then implies that each agent $i \neq 2$ obtains b with probability $\frac{1}{n-1}$ under $\rho(\succsim_2^\circ, \succsim_{-2})$.

Finally consider $\rho(\succsim_1', \succsim_2^\circ, \succsim_{-\{1,2\}})$. Since $\succsim_1'$ and $\succsim_1$ both rank a and b at the top agent 1 must, by strategy-proofness, get a or b under $\rho(\succsim_1', \succsim_2^\circ, \succsim_{-\{1,2\}})$ with the same probability as under $\rho(\succsim_2^\circ, \succsim_{-2})$. Since $\succsim_2^\circ$ and $\succsim_2$ both rank a at the top agent 2 must, by strategy-proofness, get a under $\rho(\succsim_1', \succsim_2^\circ, \succsim_{-\{1,2\}})$ with the same probability as under $\rho(\succsim_1', \succsim_{-1})$. In sum we obtain

$$\rho_1(\succsim_1', \succsim_2^\circ, \succsim_{-\{1,2\}})(a) + \rho_1(\succsim_1', \succsim_2^\circ, \succsim_{-\{1,2\}})(b) = \frac{1}{n} + \frac{1}{n-1}$$
$$\text{and } \rho_2(\succsim_1', \succsim_2^\circ, \succsim_{-\{1,2\}})(a) = \frac{1}{n-1}.$$

So when agents 1 and 2 declare $\succsim_1'$ and $\succsim_2^\circ$ at $\succsim$ agent 1's probability to receive one of his two most preferred houses and agent 2's probability to receive his most preferred house respectively increase from $\frac{2}{n}$ to $\frac{1}{n} + \frac{1}{n-1}$ and from $\frac{1}{n}$ to $\frac{1}{n-1}$. Consequently, ρ is not group-strategyproof. □

The proof goes through unchanged if we only consider matching problems without outside options $(\hat{\Omega}, \hat{X})$. For $x \in X$ to be Pareto optimal at some $\succsim \in \hat{\Omega}$, x must match as many agents as possible. So if $\rho : \hat{\Omega} \to \Delta X$ is ex-post Pareto optimal, then any x in the support of some $\rho(\succsim)$ is a matching without outside options ($x \in \hat{X}$) and we obtain the following corollary.

Corollary 1. *If there are at least three agents and three houses, then no ex-post Pareto optimal, group-strategyproof mechanism $\rho : \hat{\Omega} \to \Delta \hat{X}$ for matching problems without outside options treats equals equally.*

Theorem 1 and Corollary 1 stand in the tradition of impossibility results on fair, efficient, and non-manipulable mechanisms for matching problems. As already discussed in the introduction, Bogomolnaia and Moulin's (2001) impossibility result imposes a weaker notion of non-manipulability (strategy-proofness), a stronger notion of efficiency (ordinal efficiency), and the same criterion of fairness (equal treatment of equals). Nesterov (2014) combines the weaker notions of non-manipulability and efficiency discussed here with a stronger criterion of fairness to show that no envy-free mechanism is ex-post Pareto optimal and strategyproof.

4. Non-bossiness under randomization

Following Satterthwaite and Sonnenschein (1981) $\rho : \tilde{\Omega} \to \Delta X$ is **non-bossy** if $\rho_i(\succsim) = \rho_i(\succsim_i', \succsim_{-i}) \Rightarrow \rho(\succsim) = \rho(\succsim_i', \succsim_{-i})$ holds for all triples $(i, \succsim, \succsim_i')$. Fix a set of mechanisms $M = \{\rho^1, \rho^2, \cdots \rho^K\}$ with $\rho^k : \tilde{\Omega} \to \Delta X$ for all $1 \leq k \leq K$. Then the mechanism $\rho^* : \tilde{\Omega} \to \Delta X$ **arises out of a randomization** over M if there exists a lottery π on $\{1, \cdots, K\}$ with $\pi(k) > 0$ for all $1 \leq k \leq K$ and

$$\rho^*(\succsim)(x) = \sum_{1=k}^{K} \pi(k)\rho^k(\succsim)(x) \text{ for all } x \in X.$$

If π is the uniform distribution on set M, then ρ^* **arises out of uniform randomization** over M.

Theorem 2. *Let $\rho^* : \tilde{\Omega} \to \Delta X$ arise out of a randomization over $\{\rho^1, \cdots, \rho^K\}$ where each $\rho^k : \tilde{\Omega} \to \Delta X$ is strategyproof and non-bossy. Then ρ^* is non-bossy.*

Proof. Fix $(i, \succsim, \succsim_i')$ such that $\rho_i^*(\succsim) = \rho_i^*(\succsim_i', \succsim_{-i})$. Suppose $\rho_i^{k^*}(\succsim) \neq \rho_i^{k^*}(\succsim_i', \succsim_{-i})$ held for some $1 \leq k^* \leq K$. For h^*, the $\succsim_i$-best house to which the two lotteries assign different probabilities, we have $\sum_{h \succsim_i h^*} \rho_i^{k^*}(\succsim)(h) \neq \sum_{h \succsim_i h^*} \rho_i^{k^*}(\succsim_i', \succsim_{-i})(h)$. Since all ρ^k are strategyproof, $\sum_{h \succsim_i h^*} \rho_i^k(\succsim)(h) \geq \sum_{h \succsim_i h^*} \rho_i^k(\succsim_i', \succsim_{-i})(h)$ holds for all $1 \leq k \leq K$. In combination with $\pi(k^*) > 0$ (as required by ρ^* arising out of a randomization over $\{\rho^1, \cdots, \rho^K\}$) we obtain

$$\sum_{h \succsim_i h^*} \rho_i^{k^*}(\succsim)(h) > \sum_{h \succsim_i h^*} \rho_i^{k^*}(\succsim_i', \succsim_{-i})(h) \Rightarrow$$

$$\sum_{1=k}^{K} \pi(k) \sum_{h \succsim_i h^*} \rho_i^k(\succsim)(h) > \sum_{1=k}^{K} \pi(k) \sum_{h \succsim_i h^*} \rho_i^k(\succsim_i', \succsim_{-i})(h) \Rightarrow$$

$$\sum_{h \succsim_i h^*} \rho_i^*(\succsim)(h) > \sum_{h \succsim_i h^*} \rho_i^*(\succsim_i', \succsim_{-i})(h),$$

a contradiction to the assumption that $\rho_i^*(\succsim) = \rho_i^*(\succsim_i', \succsim_{-i})$. So $\rho_i^k(\succsim) = \rho_i^k(\succsim_i', \succsim_{-i})$ holds for all $1 \leq k \leq K$. Since every ρ^k is non-bossy, $\rho^k(\succsim) = \rho^k(\succsim_i', \succsim_{-i})$ holds for all $1 \leq k \leq K$, implying $\rho^*(\succsim) = \rho^*(\succsim_i', \succsim_{-i})$. In sum, ρ^* is non-bossy. □

Since Theorem 2 holds for the arbitrary domain $\tilde{\Omega}$, it holds in particular for Ω and $\hat{\Omega}$, housing problems with and without outside options. Theorem 2 also applies if we replace the set of all matchings X with any other discrete space of allocations Y. To see that strategy-proofness cannot be dropped from Theorem 2 consider a matching problem with $n = 3$ and $H = \{a, b, c\}$ and three (deterministic) mechanisms α, β and γ, defined by the following table (where $\succsim_1^*$ is such that $a \succ_1^* b \succ_1^* c$):

	$\alpha(\succsim)$	$\beta(\succsim)$	$\gamma(\succsim)$
$\succsim_1 = \succsim_1^*$	(a, c, b)	(b, c, a)	(c, b, a)
$\succsim_1 \neq \succsim_1^*$	(b, a, c)	(c, a, b)	(a, c, b).

Agent 1 alone determines the matching in these mechanisms. Since agent 1's match under α, β or γ changes if and only if the others' matches change and α, β, and γ are non-bossy. However the mechanism $\rho^\circ : \Omega \to \Delta X$ that arises out of the uniform randomization over α, β and γ is bossy. For any $\succsim$, agent 1 faces the uniform lottery over $\{a, b, c\}$ under $\rho^\circ(\succsim)$. However, the other agents' lotteries over houses depend on agent 1's announcement. Agent 2, for example, never gets house a if agent 1 announces $\succsim_1^*$, but gets a with probability $\frac{2}{3}$ if agent 1 announces any other preference. In a similar vein, non-bossy mechanisms might arise out of the randomization over bossy mechanisms. Theorem 2 does not extend to a larger domain of preferences that allows for indifferences. Examples to prove the latter two claims are available on request.

5. Random serial dictatorship

Papai (2000) showed that a deterministic mechanism is group-strategyproof if and only if it is strategyproof and non-bossy. I use random serial dictatorship as an example to show that this equivalence does not hold for random mechanisms. For any permutation $p : N \to N$ define the serial dictatorship sd^p as the deterministic mechanism that uses p to sequentially entitle agents to choose houses. So $sd^p_{p(1)}(\succsim)$ is agent $p(1)$'s most preferred house according to $\succsim_{p(1)}$, $sd^p_{p(2)}(\succsim)$ is the $\succsim_{p(2)}$-preferred house among all remaining ones and so forth. Random serial dictatorship $rsd : \hat{\Omega} \to \Delta\hat{X}$ arises out of a uniform randomization over all serial dictatorships $\{sd^p \mid p : N \to N$ a permutation $\}$. Any sd^p is strategyproof, non-bossy, and Pareto optimal. Random serial dictatorship treats equals equally. Theorems 1 and 2 therefore imply that random serial dictatorship, which is known to be strategyproof, is non-bossy but not group-strategyproof.

References

Barbera, S., Berga, D., Moreno, B., 2014. Group strategy-proofness in private good economies. UAB, Barcelona. Mimeo.

Bogomolnaia, A., Moulin, H., 2001. A new solution to the random assignment problem. Journal of Economic Theory 100, 295–328.

Erdil, A., 2014. Strategy-proof stochastic assignment. Journal of Economic Theory 151, 146–162.

Kesten, O., Kurino, M., 2013. Do outside options matter in school choice? A new perspective on the efficiency vs. strategy-proofness trade-off. Carnegie Mellon University. Mimeo.

Nesterov, A., 2014. Fairness and Efficiency in a Random Assignment: Three Impossibility Results. WZB, Berlin. Mimeo.

Papai, S., 2000. Strategyproof assignment by hierarchical exchange. Econometrica 68, 1403–1433.

Satterthwaite, M., Sonnenschein, H., 1981. Strategy-proof allocation mechanisms at differentiable points. Review of Economic Studies 48, 587–597.

Svensson, L.G., 1999. Strategy-proof allocation of indivisible goods. Social Choice and Welfare 16, 557–567.

Thompson, W., 2016. Non-bossiness. University of Rochester. Mimeo.